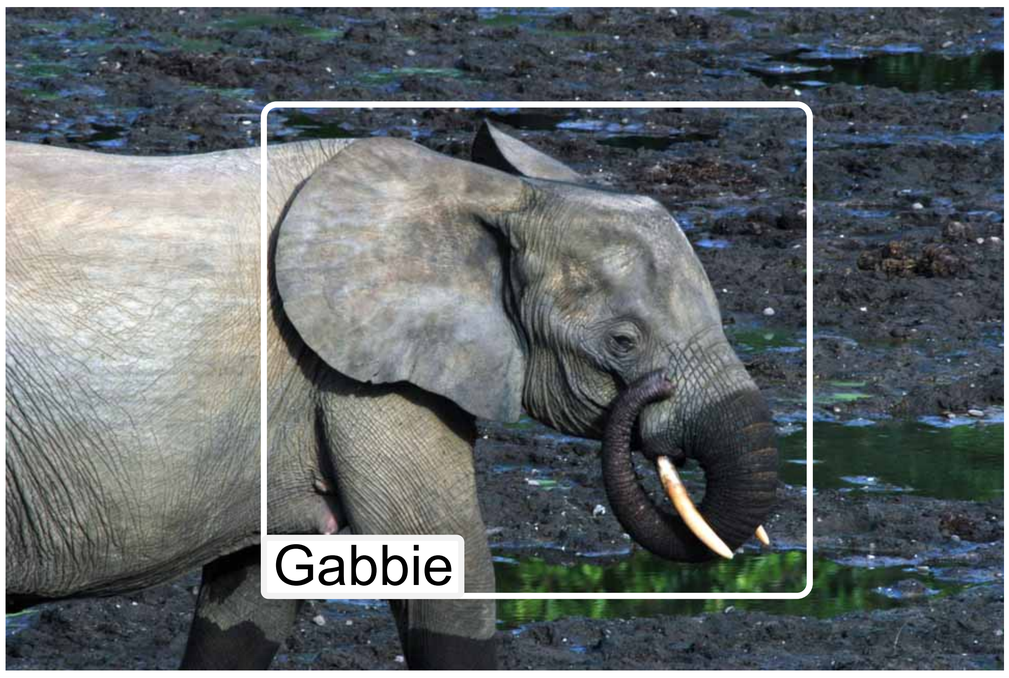
Towards Automatic Identification of Elephants in the Wild

Matthias Körschens, Björn Barz, Joachim Denzler

Computer Vision Group, Friedrich Schiller University Jena, Germany

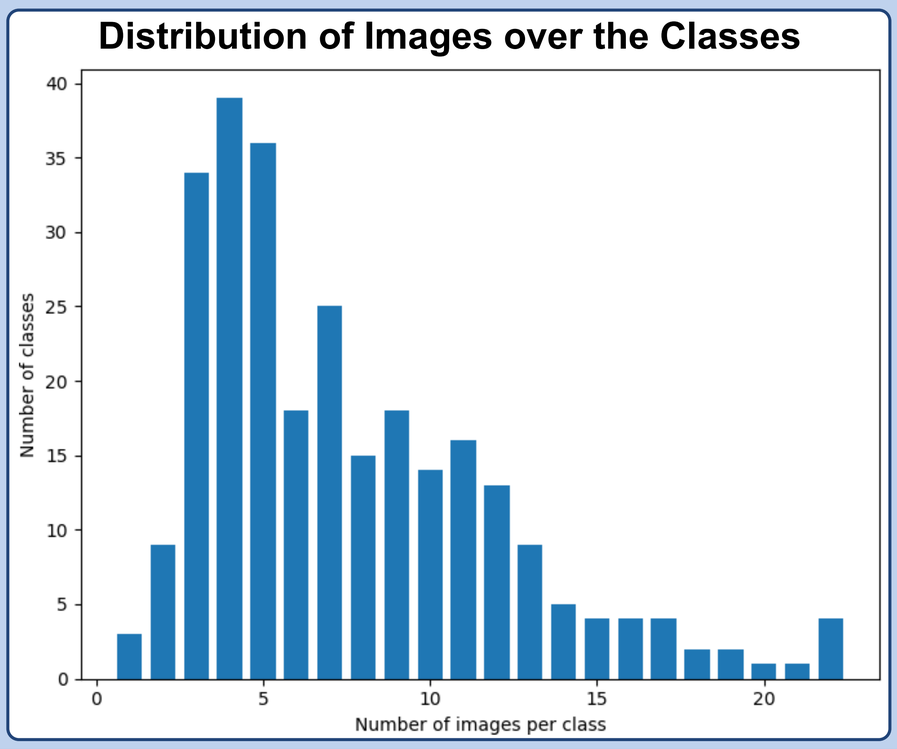
Motivation

- Supportive system for identification of elephants in the Kongo needed
- Cooperation project with biologists from the Elephant Listening Project (ELP)[1] of the Cornell University
- Many individuals (~4000 in the last 25 years)
- Up until now: identification via record cards done by a single person
- Distinctive features archived, but finding the correct elephant is exhausting
- Similar appearance with only subtle differences
- System should list a number of most probable individuals, which can then be easily distinguished by the user

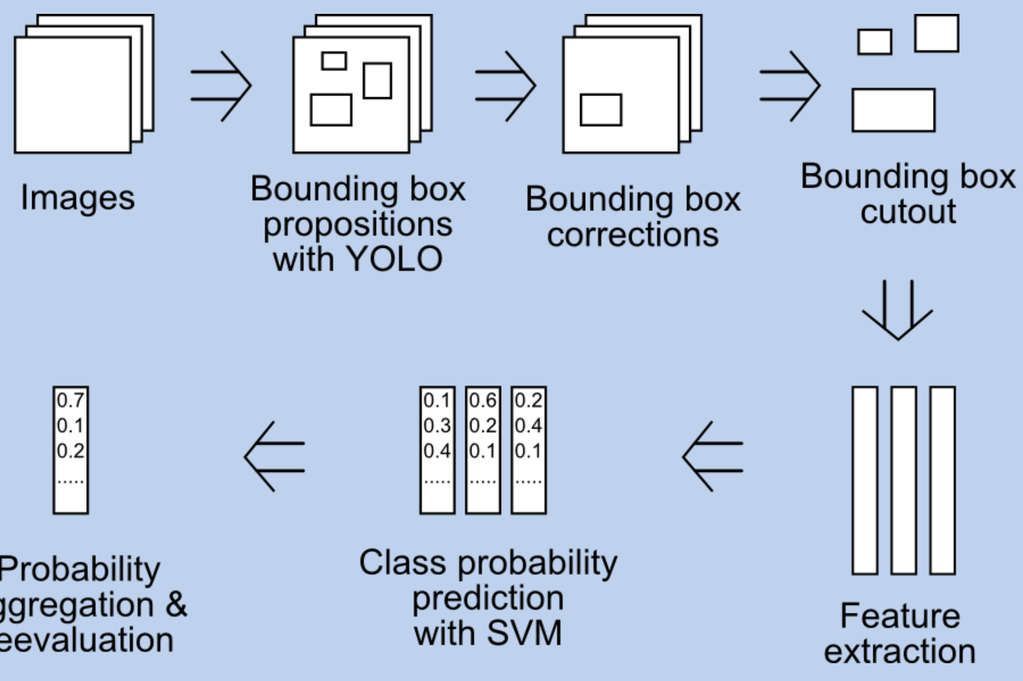


The Dataset

- 276 classes, in total 2078 images
- Images taken manually in the field over many years
- Big challenge: unbalanced distribution of images over the classes
- 75% split: 1573 training and 505 test images
- Because of unbalanced distribution: no independent k-fold split possible

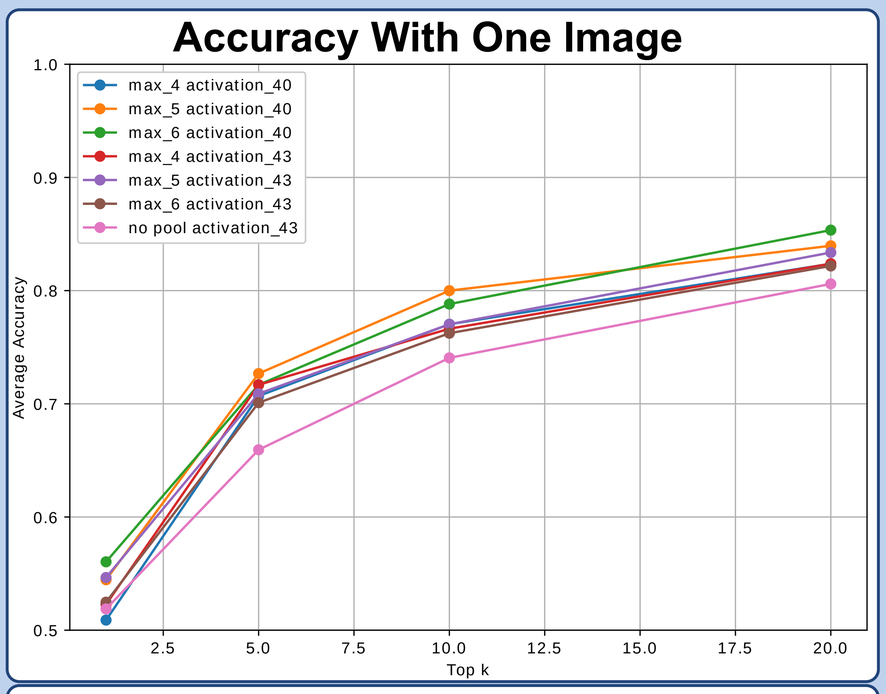


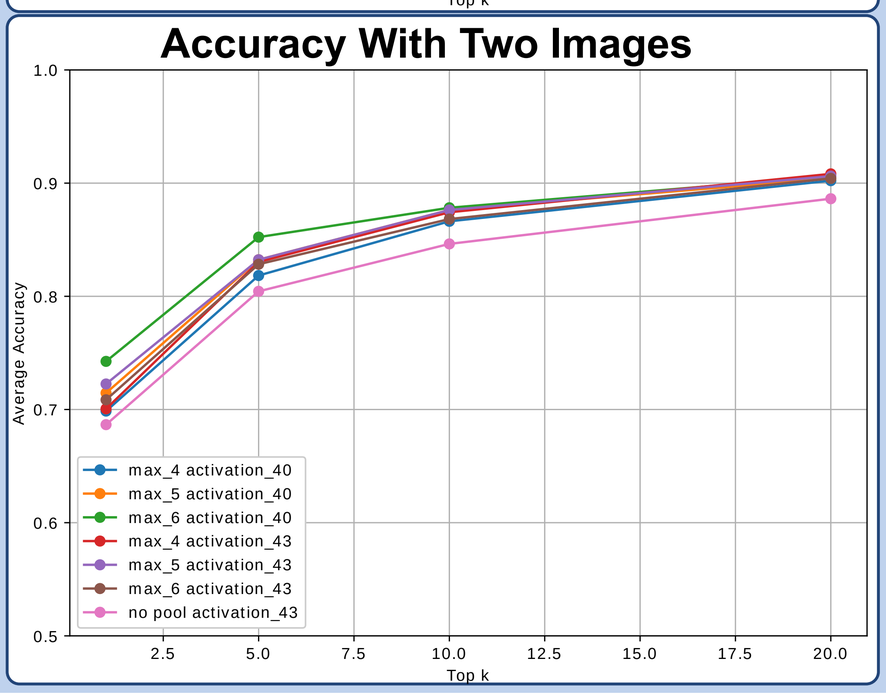
Identification Process



Experimental Results

- Measure of classification accuracy using one and two images
- Comparison of different pooling sizes (max_*n*)
- Comparison of different extraction layers (activation_*n*)





Conclusion

- Using one image on held-out test set:
  - Top 1 accuracy of up to 56%, per-class: 49%
  - Top 10 accuracy of up to 80%, per-class: 74%
- Using two images on held-out test set:
  - Top 1 accuracy of up to 74%, per-class: 59%
  - Top 10 accuracy of up to 88%, per-class: 79%
- Earlier CNN extraction layers can generate better results
- Better results possible with well-balanced dataset
- Usage as interactive system possible (realized in this project)

System Details

- Extraction of head bounding boxes with YOLO[2]
- Data augmentation: Flipped images for training
- Features extracted from Convolutional Neural Network (ResNet50[3]) pretrained on ImageNet[4]
- Extraction from earlier layers instead of last
- Max pooling of extracted features
- Classification with linear Support Vector Machine (SVM)[5]

References

[1] http://www.birds.cornell.edu/brp/elephant/

[2] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pages 779–788, 2016.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pages 770–778, 2016.

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009.

[5] Corinna Cortes and Vladimir Vapnik. Support-vector networks. Machine learning, 20(3):273–297, 1995.